DOCUMENT RESUME

ED 409 187 SE 060 348

AUTHOR Henderson, Ronald W.; St. John, Linda

TITLE Thematically Integrated Middle School Mathematics: A School-University-Business Partnership.

INSTITUTION California Univ., Santa Cruz.
SPONS AGENCY California Academic Partnership Program.; Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Apr 97
NOTE 62p.
CONTRACT R117G10022
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; At Risk Persons; *Elementary School Mathematics; Group Dynamics; *Hispanic Americans; Junior High Schools; Manufacturing Industry; Mathematics Instruction; Middle Schools; *Partnerships in Education; Population Trends; Problem Solving; *Relevance (Education); Small Group Instruction; Social Background; Student Attitudes; *Thematic Approach
IDENTIFIERS *Hispanic American Students; Latinos; Middle School Students

ABSTRACT

This report examines the accomplishments of a collaborative project to enhance student learning and motivation in mathematics in a middle school serving a changing population of students, the majority of whom are Latino. The partnership joined together the mathematics teachers and site administrator of the school, university mathematicians and educational researchers, and the management and workers from a research-and-development-oriented electronics firm. The partners in the project believed that a thematic approach to mathematics instruction might address a number of the hypothesized causes of low achievement. Results suggest that the project school is doing an especially effective job at the seventh grade level. All students showed improvement against the national norms regardless of mathematics curriculum. Eighth grade data were less encouraging but the students who were taught using a thematic curriculum held their own against the national norms. Students in the thematic class scored better on the Mathematics in Manufacturing criterion test that sampled material covered in a manufacturing unit. Contains 33 references. (PVD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 409 187

Thematically Integrated Middle School Mathematics:

A School-University-Business Partnership

Ronald W. Henderson and Linda St. John[1]

University of California, Santa Cruz

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

R.W. Henderson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

April, 1997

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Project Co-Directors
Edward M. Landesman
Margaret White

Project Co-Principal Investigators
Edward M. Landesman
Ronald W. Henderson

The work reported here was supported in part by funds from the California Academic Partnership Program for the project entitled "Thematic Mathematics in the Middle School" (Project 41), and in part by support from the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education, under Cooperative Agreement No. R117G10022, through a project of the National Center for Research on Cultural Diversity and Second Language Learning, entitled "The Role of Sociocultural, Instructional and Motivational Factors in the Development of Higher Order Cognitive Processes in Mathematics Among Language Minority Students".

BEST COPY AVAILABLE

SE 060 318

Thematically Integrated Middle School Mathematics:

A School-University-Business Partnership

This report examines the progress and accomplishments of a collaborative project implemented at Washington Middle School, Salinas Unified High School District. This partnership joined together in an academic partnership the mathematics teachers and site administrator of a middle school serving a changing population of students (the majority of whom are now Latino), university mathematicians and educational researchers, and the management and workers of a research and development oriented electronics firm. The general aim of the partnership is to enhance students' learning and motivation in mathematics.

Latino students, whose mathematics learning and future participation in mathematics are the central concern of this project, comprise a rapidly growing portion of the population of the United States. Whereas many of these students do well in school, a disproportionate number of them are considered at risk for educational failure and underachievement (Haycock & Navarro, 1988; Henderson & Landesman, 1995; MacCorquodale, 1988; Policy Analysis for California Education, 1990, Valencia, 1991). Low achievement among this rapidly growing segment of the student population is a matter of grave concern to educators and policy makers and a source of frustration to those countless numbers of teachers who put forth their best effort, yet fail to see their students make satisfactory educational progress. Accounts in the popular press of lagging achievement

among Latino students both reflect and promote the widely held public sentiment that, despite substantial expenditures of money, the public schools have not succeeded in their attempts to raise the achievement of students in this burgeoning population (Chavez, 1993; LaFranchi, 1984; Mendoza, 1995).

The present generation of Latino students is California's fastest growing population. Although the popular media tends to focus attention on demographic changes in urban environments, a large portion of the burgeoning Latino population makes its home in small to mid-sized communities of California's agricultural heartland where educators are faced with the major challenge of preparing today's students to assume productive roles in the workplaces of tomorrow. A large proportion of Latino workers are employed in agriculture and agriculture-related industries. Even if their employment opportunities were not so restricted, the kinds of jobs in which former generations found some degree of security and upward social mobility are being exported to developing countries (e.g., Kamm, 1994). Future occupations will require more education than those of the recent past, and a major portion of the more satisfying and remunerative careers will require competence in mathematics. Within the United States, employment in manufacturing and related industry will decrease as we become increasingly dependent on workers providing what former US Secretary of Labor Robert B. Reich (1991) called "symbolic-analytical services." These services encompass a wide range of occupations which have in common the need for college preparation in mathematics. Increasingly for Latinos as for all Americans, job-oriented educational opportunities will hinge on certain basic prerequisites such as algebra, geometry, and pre-calculus, even though the ways

these topics are presented are themselves subject to change.

The nation's economic welfare depends upon the development of a population educated to engage in symbolic analysis, but so too does the welfare of individuals. The kinds of economic shifts forecast by Reich and others became salient for mathematics educators over a decade ago (Romberg & Zarinnia, 1987; Zarinnia & Romberg, 1987), stimulating a call for reform in the nature of school mathematics in this country. The resultant mathematics reform effort calls for a shift in emphasis from a "transmission of knowledge" model, which emphasizes computation, to approaches that stress the "conceptual, analytical, and problem-solving techniques of mathematics" (Sowder, 1989). This kind of curricular reform necessitates a shift away from the decontextualized atmosphere of traditional mathematics teaching and learning in classrooms, and toward a contextualized approach in which there are explicit real world connections for mathematical ideas. This shift also implies language-intensive teaching that affords students rich opportunities to solve and form mathematics problems together in shared contexts. On a theoretical level, this shift is consistent with the view that cognitive development takes place within a social context and is, itself, a product of social interaction (Tharp & Gallimore, 1988; Vygotsky, 1962, 1978). These ideas were central to the proposition that a Thematic approach to mathematics might be especially effective for Latino students. Although there is, as yet, little supportive empirical evidence on the effectiveness of contextualized mathematics instruction (of which thematic instruction is one form), the partners in the present project were of the opinion that this contextualized approach could address a number of the hypothesized causes of low achievement.

Previous reports evaluated the first (Henderson, 1994) and second (Henderson, St. John, & Youpa, 1994; Henderson, 1995) years of the implementation phase of this project. Those results were generally encouraging, showing positive growth in mathematics learning among seventh graders in all three different treatment groups that were implemented in those two years. The present report provides an evaluation of the effects of this program on mathematics achievement, student self-perceptions, and attitudes relating to mathematics[1] during the third year of program implementation,[2] summarizes the overall results for the life of the project, and comments on lessons learned from the project. The evaluation effort involved on-going participant observation that contributes to the qualitative analysis of the process of implementation. These analyses provide formative evaluation information that contributes to the program and are described briefly in the present report.

The design of the evaluation called for entering seventh grade students at Washington Middle School to be assigned randomly to one of three treatment conditions. The experimental condition was a thematic approach to mathematics instruction, with stu-

[1]The program development and implementation aspects of this project and some operational costs are supported by the California Academic Partnership Program. Research and evaluation activities are largely activities of a project of the National Center for Research on Cultural Diversity and Second Language Learning, University of California, Santa Cruz. That project, entitled " The Role of Sociocultural, Instructional and Motivational Factors in the Development of Higher Order Cognitive Processes in Mathematics Among Language Minority Students," addresses the problem of generally low levels of mathematics achievement and participation among Latino populations. The aim of this research and its associated curricular reform activities is to gain an understanding of the sociocultural, instructional, and motivational dynamics that influence learning outcomes in mathematics among Latino students.

[2]Implementation was preceded by a year of preliminary work under a planning grant from the California Academic Partnership Program.

dents being assigned to classrooms on a heterogeneous basis. A traditional approach in which students were assigned to homogeneous classes on the basis of prior mathematics achievement, and in which instruction was rather tightly bound to the textbook, constituted the second treatment condition. Students in the third treatment were assigned to heterogeneously (by mathematics achievement) constituted classrooms and received a "Blend" of textbook-based and project-based instructional approaches.

These same three treatments were in effect during all three years of implementation, with variations to be described later in this report. During the first year of the project, these programmatic alternatives were introduced at the 7th grade level. These treatments were extended to include grade eight during years two and three. The evaluation, then, examined the mathematics achievement and school-related affective responses of students assigned to each of these three conditions; Thematic, Traditional, and Blend. The original intent was for students who were assigned to a given instructional treatment at the beginning of grade 7 to continue in that same treatment during 8th grade. Thus, the cumulative effects of participation in a given instructional approach could be assessed. Unfortunately, the realities of middle school scheduling prevented the systematic implementation of this strategy.

Method

Subjects

The students whose learning and school-relevant affect were examined for the first time in this evaluation were all seventh and eighth graders attending Washington Middle School during the 1995-96 academic year. Analyses were based on all students

for whom the relevant measures of a given analysis were available. Most of the analyses examine pretest to posttest performance across the three treatment groups at Washington Middle School. In addition, the pre- to posttest achievement of the 1995-96 cohort of 7th and 8th graders (cohort 3) was compared with cohorts from the previous years.

In the analysis of the first year of implementation we also compared the mathematics achievement of seventh graders at Washington Middle School with that of two other middle schools in the same school district. Those comparisons were not pursued in the present report because the populations served by these different schools are not sufficiently comparable in terms of ethnic distribution and the proportion of students with limited English proficiency (LEP).

Within the school in which the project was conducted, the student population is divided into families comprised of 150 students. This arrangement is intended to ease the transition to middle school by avoiding the impersonal character of traditional departmentalized structures. With minor exceptions, the students in a given family spend their entire schedules with just four teachers. As was the case in the first two years of implementation, this arrangement placed restrictions on the degree to which random assignment of students to treatment groups could be implemented because, for scheduling purposes, the students assigned to the school's Gifted and Talented Enrichment (GATE) program were all assigned to a single family unit. It was necessary to make a number of adjustments to accommodate students for whom the options for class assignment were restricted because of their participation in groups such as the GATE program. For example, while there were no special mathematics classes specifically arranged for students

who participated in the GATE program, it was necessary to assign students who were selected for this program to classes designated for GATE students in other subjects, thus reducing the "degrees of freedom" available for the assignment of these students to instructional conditions. Within the Traditional treatment group, assignment to classes was based on achievement level, as had been the custom at this school prior to the CAPP intervention. Most GATE students were assigned to the Traditional treatment group, with a small number scattered throughout classes in the Thematic condition. This resulted in substantial differences in the entry level mathematics achievement of students in the different instructional conditions.

Just as the Thematic instructional approach for seventh grade was under development during the initial year of the project (see Henderson, 1994), thematic curricula and instructional practices for the eighth grade were being developed even as the program was being implemented during year 2. Refinements were still being worked out in the Thematic curricula for both grades during year 3. The major Thematic activity for grade eight was a thematic unit called *Mathematics in Manufacturing*, developed as a product of the three-way collaboration among the school, the university, and the business partner.

Measures

Achievement. The school district administers selected subscales from the Comprehensive Test of Basic Skills (CTBS) during the spring and fall terms of each academic year. The CTBS *Mathematics Concepts and Applications* subtest was used to assess overall mathematics achievement. For seventh graders testing results from fall were used

as pretest scores, whereas scores from the previous spring were used as the pretest measure for eighth graders because achievement tests are not administered to continuing eighth graders in the fall. The *Concepts and Applications* subtest, together with other subtests of the CTBS, was administered as a posttest in early March, 1995. During the second year of the project a member of the teaching staff expressed concern that administration of the test in March would be too early (and on the heels of a flood that displaced families and disrupted the schooling process) to assess adequately the students' growth in mathematics achievement. Therefore, she arranged for a repeat administration of the Concepts and Applications subtest in June. The posttest measures used in the analyses reported for the second year were based on the June administration (Henderson, St. John & Youpa, 1994). Since the June testing took place at the end of the testing window designated by the test publisher, the posttest results could have been slightly inflated. Both March and June testing results are reported for the third and final year of the project.

A locally designed test was also administered on a pre- and posttest basis to seventh graders at Washington Middle School. This test assessed specific skills and concepts that were included in a curriculum unit (*Quincy Market*) that was adapted for use in the Thematic treatment group, but that are also standard elements in the school's traditional grade seven mathematics curriculum (e.g., measurement, ratio and proportion, perimeter and area, relationships among the properties of circles, and percents). Both multiple choice and performance-based approaches to assessment were incorporated into the Quincy Market measure that was used during the second year of the project. The performance-based assessment was based on work done as a part of master's degree projects

(Moss, 1995; O'Donnell, 1995). Lacking the additional assistance provided by these two graduate students, it was not possible to use the performance based assessment items in the third year.

During both year two and year three, progress on skills and concepts that were included in the eighth grade *Mathematics in Manufacturing* thematic curriculum were assessed at the beginning and end of the instructional unit. The Quincy Market pre- and posttesting for the Thematic group was accomplished at the beginning and end of the Quincy Market instructional unit; a period of approximately six weeks. For other treatment groups, the posttest was administered at the end of the year because the skills and assessed by the test were covered at different times during the academic year in the other treatments.

Self-Perception. A Self-Perception instrument entitled "How I See Myself" was administered at the beginning and end of each year. This instrument included 5 subscales, three of which (self-perceptions of Academic Competence, Social Acceptance, and Global Self-Worth) were adapted from Susan Harter's (1985) Adolescent Self-Perception Profile. The remaining two scales, Challenge-Seeking/Persistence and Goal Orientation (Henderson, 1991) assessed variables thought to be involved in the development of mathematics motivation. These measures will be used in on-going studies of motivational development, but for present purposes we hypothesized that the Thematic treatment would have a favorable effect on Goal Orientations and Challenge-Seeking/Persistence because of its emphasis on the integrated application of mathematics skills and concepts to ongoing, multifaceted problems. Goal orientations have been

identified as potentially important contributors to academic achievement in general (Dweck, 1986; Dweck & Leggett, 1988; Elliot & Dweck, 1988) and in mathematics, in particular (Peterson, 1988), but little is known about the dynamics of goal orientations in natural classroom settings or for different ethnic groups. The overriding goal of performance-oriented students is to gain positive judgments and to avoid negative judgments of their competence by others. In contrast, students displaying a learning orientation seek to increase their competence, regardless of whether their confidence in their ability to succeed at the task is high or low. Learning-oriented students are more likely than their performance-oriented peers to explore, take initiative, and pursue tasks that promote intellectual growth, including an inclination to seek intellectual challenge and to persist in the face of difficulty. Therefore, we expected Challenge Seeking/ Persistence and goals characterized by a learning orientation to predict learning outcomes. Our reasoning was that working on long range problems requiring higher order thinking processes would influence students in the Theme treatment to be more concerned with learning than with performance outcomes, and that thematic instruction would have a positive influence on students' willingness to persist in the face of difficult challenges.

Mathematics Attitudes Questionnaire. The attitudes and beliefs that students hold toward mathematics may play an important role in their mathematics achievement, and in their willingness and interest in participating in mathematics (Haladyna, Shaughnessy & Shaughnessy, 1983; McLeod, 1985; Silver, 1985). Some educators have argued that disinterest and relatively poor mathematics achievement among students of Mexican descent is the result of early and repeated failure in school, and that these experiences

lead to poor academic self-concept, negative attitudes toward school subjects, and alienation from school. Others have suggested that Hispanic children are directed away from mathematics as the result of negative cultural stereotypes of mathematicians (Cocking and Chipman, 1988). These beliefs have not been widely confirmed, and some research with other samples from the population involved in the present study has found quite positive attitudes toward mathematics and relatively high levels of mathematics self-concept (Henderson & Landesman, 1995). Therefore, the present evaluation examined attitudes, with the long range aim of understanding the connection between mathematics attitudes and achievement, and how that relationship might be affected by variations in the approach to mathematics instruction.

Mathematics attitudes were assessed with a questionnaire utilizing a 5-point likert-like format, with responses ranging from strongly agree to strongly disagree. The instrument included six subscales, the items of which were randomly distributed throughout the questionnaire. The subscales are Mathematics Self-Concept, Mathematician Stereotypes, Mathematics Attitudes, Mathematics Futures, Nature of Mathematics, and Parental Support and Encouragement in Mathematics, which includes parallel sets of items referencing mothers and fathers. The subscales relating to parental encouragement and support were administered for use in subsequent analyses and the results are not presented in the present report. In addition to the subscales, an item assessing students' perceptions of mathematics as a more appropriate field of study for boys than for girls was included in the analysis. This measure is referred to as Gender Stereotype.

The Mathematics Self-Concept scale consisted of 5 items such as "I am good at

working with numbers," and "I am good at mathematics," intended to assess students' confidence in their ability to learn mathematics. The Math Attitudes subscale (8 items) employed items such as "Mathematics is boring to me," "Mathematics is interesting," and "I really enjoy math class" to assess students' interest in and enjoyment of mathematics.

The Mathematics Future subscale (5 items) was included to assess interest in working at an occupation that uses math and expectations that their work as adults would require mathematics, as indicated by responses to items such as "When I am older, I expect to work in an area that requires mathematics," and "I would like to have a job working with math when I finish school." These items were adapted from items used in the National Assessment of Educational Progress (Dossey, Mullis, Lindquist, & Chambers, 1988).

Three items were adapted from the National Assessment of Educational Progress (Dossey, Mullis, Lindquist, & Chambers, 1988) to assess students' conceptions of the nature of mathematics. These items asked students to respond to statements such as "Learning mathematics is mostly memorizing," and "There is always a rule to follow in mathematics."

The Stereotypes of Mathematicians subscale (6 items) was included to examine the assertion that Mexican-Americans attribute characteristics to mathematicians which are contrary to their cultural values (Cocking & Chipman, 1988); characteristics such as being impersonal and distant from other people and being unable to engage in normal family life. Items developed for this purpose included statements such as "Mathematicians don't show much interest in people," and "Mathematicians can have a

normal family life."

Results

Grade Seven Achievement

Figure 1 provides a comparison of 7th graders' performance on the CTBS Concepts and Applications subtest for the first two operational years of the project. All treatment groups in the year two cohort appear to have exceeded the very good performance achieved by the initial cohort during the 1993-94 academic year.

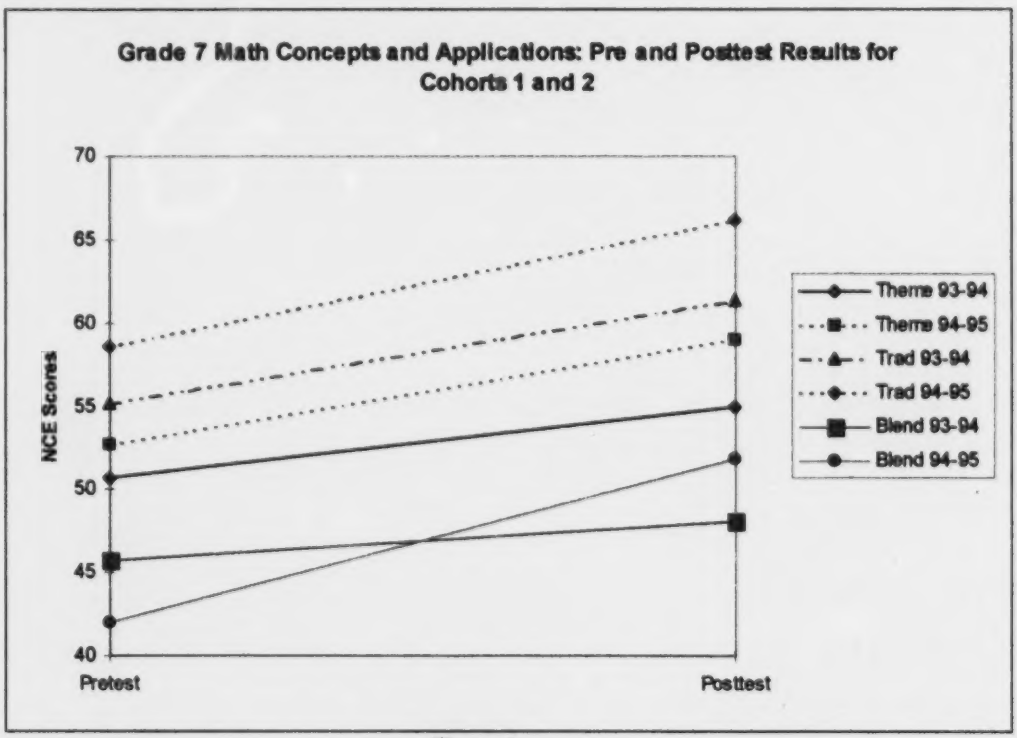

Figure 1

The results shown in Figure 1 for cohorts 1 and 2 can be compared with those of the third 7th grade cohort, depicted in Figure 2. We examined the achievement of seventh grade students who experienced the different treatment conditions by means of a 2 (trials) x 3 (treatments) x 2 (sex) repeated measures analysis of vari-

BEST COPY AVAILABLE

ance, with pre- and posttest scores from the Mathematics Concepts and Applications subscale of the CTBS as the repeated measure. The analysis revealed significant main effects for condition, F (1, 396) 31.17, $p < .001$, and trials, F (1, 396) 7.745, $p < .01$. The results indicate that the three treatment groups made comparable pre to posttest gains. Although the increase in scores appears small in absolute value terms, the changes were statistically significant.

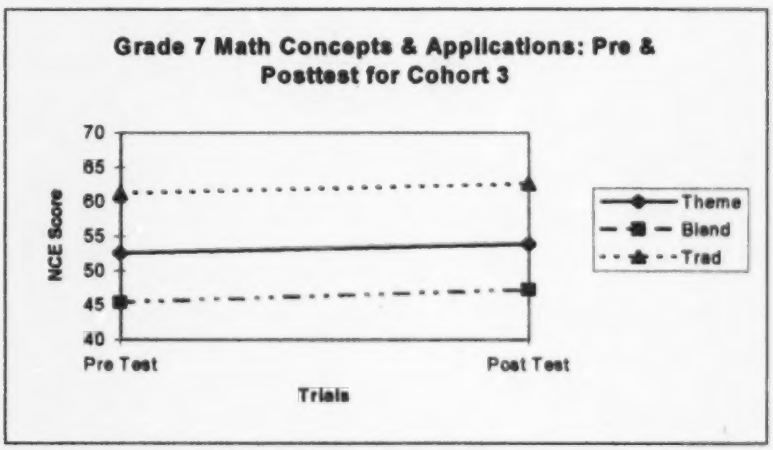

Figure 2

The groups displayed significantly different levels of mathematics achievement, as indicated by the treatment main effect. Initial differences in the achievement levels of the treatment groups are attributable to the placement practices mentioned earlier. The Blend treatment group, which had the lowest entering scores, included transitional classes comprised of students who were moving from bilingual classes to classes conducted entirely in English. The lack of a significant trials by treatment interaction demonstrates that the groups made parallel gains. It should also be kept in mind that these scores are based on norms that take the time of year into consideration. Therefore, an increase in NCE

scores represents a gain against the norming group. If scores did not change it would indicate that students did make gains in the number of items answered correctly, but that their position within the norm group did not change. Another consideration is that the posttests were administered in early March, at the very beginning of the norming window for these students' grade. Therefore, the scores might well give a rather conservative view of the learning that took place over the course of the entire academic year.

Since the Thematic treatment, and to some extent the Blend treatment as well, involves an integrative approach in which it might take longer for achievement gains to be consolidated, a separate analysis was conducted using June posttest data instead of the posttest data from early March. The 2 (trials) by 2 (sex) by 3 (treatments) repeated measures ANOVA yielded significant main effects for treatment, F (2, 341) 36.40, $p <$ .001, and trials, F (1, 341) 166.30, $p <$ .001. The trials by sex interaction approached significance, F (1, 341) 2.80, $p <$ .10, indicating a tendency for girls (Mean = 62.74) to

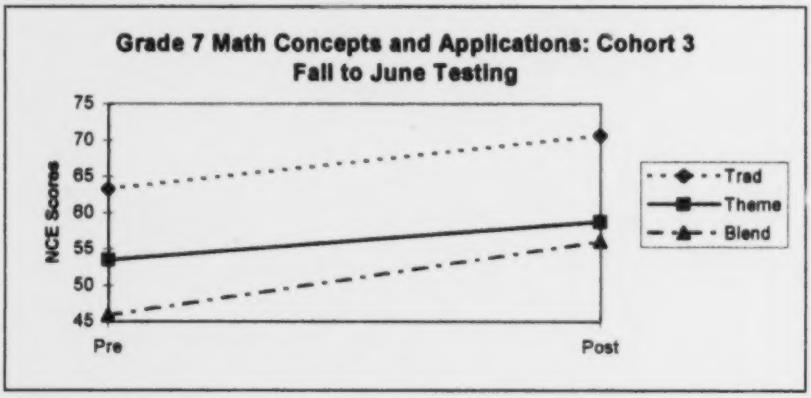

Figure 3

achieve at a slightly higher level than boys (60.86) on the posttest. Figure 3 displays the trials by treatment interaction, which was significant, $F(1, 341)\ 5.79, p < .005$.

As with the previous analysis, the treatment groups differed in both their pretest and posttest performance levels, but the later posttest data show somewhat larger changes over the course of the academic year. In addition, with the June posttest data the lines representing pre to posttest change were not parallel for the different treatment groups, as was the case with the earlier posttest data. These later posttest data show a steeper gain for the Blend treatment than for the other groups.

The remainder of the analyses of CTBS data reported here employ the June posttest data, rather than the posttest scores from tests administered in early March. Since the June testing comes at the end of the norming window, scores may reflect somewhat inflated results in relation to national norms. However, our main interest was in comparisons across the three alternative instructional treatments, rather than comparisons with national norms. A time span covering the entire academic year was judged to be most likely to reflect the entire picture of change across the school year on the kinds of tasks sampled by the CTBS Mathematics Skills and Concepts subtest.

The decision to develop and implement a thematic approach to instruction was based on an expectation that linguistic minority students would be especially likely to benefit from contextualized instruction. In order to take students' English language proficiency into account, we conducted a separate analysis using the CTBS Language Total scores as the covariate in an analysis of covariance to control statistically for differences in language achievement. The ANCOVA yielded significant main effects for treatment,

F (2, 337) 11.31, p < .001, for sex, F (1, 337) 10.35, p < .001, and for trials, F

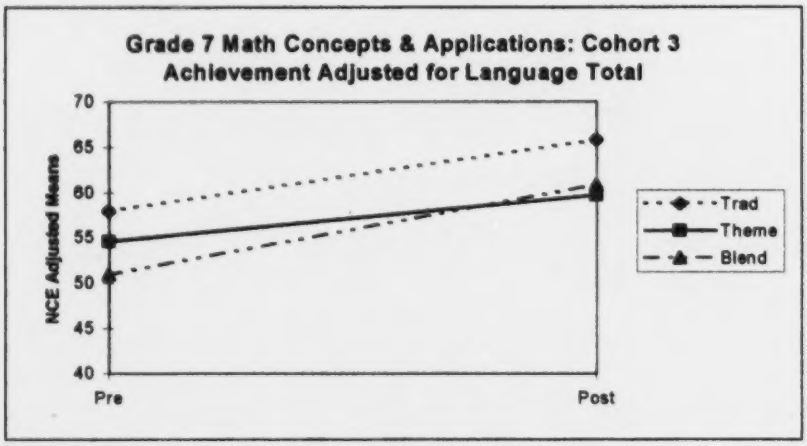

Figure 4

(1, 337) 36.97, p < .001. The treatment by trials, F (2, 337) 5.07, p < .01 and trials by sex, F (1, 337) 4.11, p < .05 interactions were also significant. Figure 4 shows the trials by treatment interaction and the sex by trials interaction is depicted in Figure 5. The trials by treatment interaction effect reflects the finding that, with statistical adjustment for differences in Language Total, the Blend group achieved the greatest pre to posttest gains. The Traditional treatment also made somewhat larger gains than did the Theme group.

Figure 5 shows that both male and female students made pre to posttest gains. With statistical adjustment for Language Total scores obtained early in the fall semester, male students displayed higher math achievement than females on both the pre and posttests, but females made greater gains than did males over the course of the academic year.

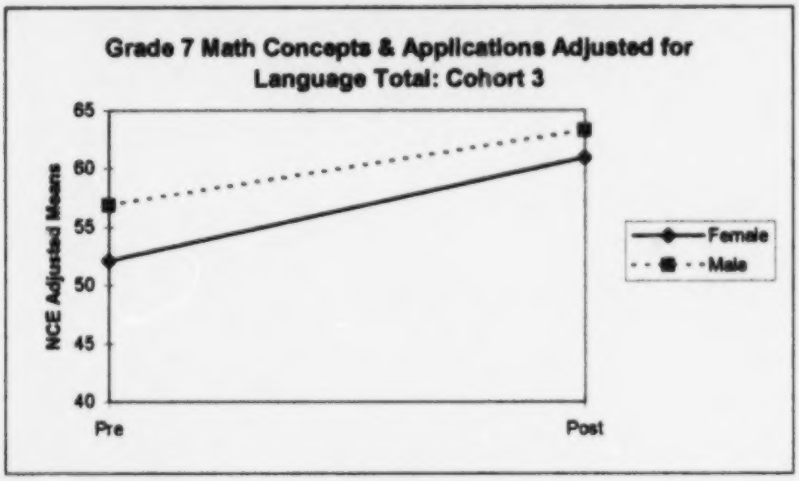

Figure 5

Unclassified Treatment: Three seventh grade classes were not included in the analyses of CTBS results presented previously. These classes were not omitted arbitrarily, but because their inclusion within any of the treatment categories would render the comparisons meaningless. Although these classes were designated to receive one or another of the three instructional treatments, both our own observations and information provided by the regular mathematics faculty at the school made it clear that the instruction received by students in these classes could not be classified according to any of the treatment categories called for by the evaluation design. These classes were subjected to a changing cast of teachers (3 over the course of the year) who were poorly qualified to teach mathematics, the hiring of whom was handled without consultation with the project school's administration or mathematics faculty, and which was judged by them to be ill-advised. In order to determine how this situation affected the learning of the students who were assigned to these classes, we conducted a separate analysis, treating these classes as a separate, "unclassified" instructional condition. For this analysis, CTBS

Mathematics Concepts and Applications NCE scores were subjected to a 2 (sex) by 4 (treatments) by 2 (trials) repeated measures ANOVA. The main effects for instructional condition (Treatment), F (3, 459) 26.903, $p < .001$, and the Trials by Sex, F (1, 459) 5.382, $p < .05$, and Trials by Condition, F 3, 459) 2.699, $p < .05$, interactions were significant. Data representing the Trials by Treatment are displayed in Figure 6.

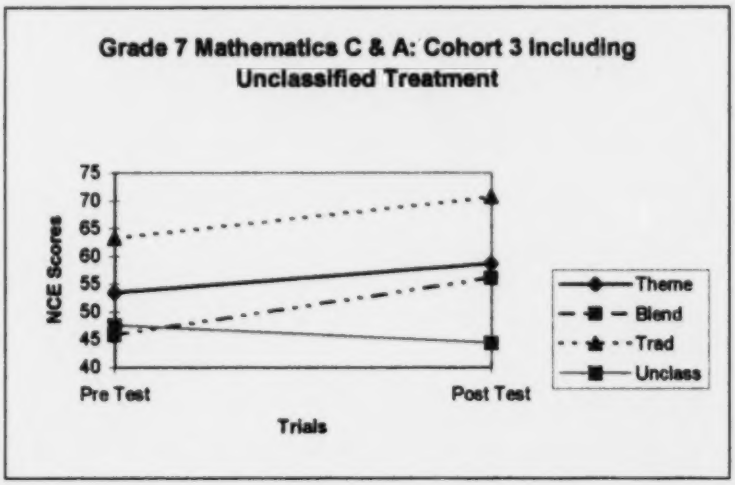

Figure 6

As shown in Figure 6, the condition by treatment effect is due to the fact that the "Unclassified" treatment lost ground against national norms during the 7th grade, whereas the other three treatments made gains. The trials by sex interaction is explained by the fact that NCE Concepts and Applications scores of 7th grade males decreased over the period of the academic year, while scores of females increased.

Quincy Market Curriculum: A major instructional activity for the seventh grade Theme group was a thematic unit entitled *Quincy Market* (Regional Math Network, 1987). Work on the Quincy Market unit began with the first year of the project and the curriculum was refined and adapted to transition into the Mathematics in Manu-

facturing unit developed for grade 8. Year 2 outcomes were assessed by means of a criterion referenced test of 20 items and by three performance-based items that were scored by means of a six point rubric. The tasks sampled by the test were all specifically within the scope of the Quincy Market theme, but the entire item set also represented material typically covered in the seventh grade mathematics curriculum at the target school. The second year results for the multiple choice portion of the test revealed significant pre to posttest gains for all treatment groups, but with a smaller relative gain by students in the traditional treatment than for the other two treatments. In order to determine if the instructional treatments had differential effects on students who differed in their mathematics achievement levels at the beginning of the school year, we conducted an additional analysis of *Quincy Market* achievement in which students were divided into high and low achievement levels (median split), based on their entry level CTBS Mathematics Concepts and Applications scores showed that the higher achieving subsets of all three treatment groups made significant gains, but among the lower achieving subsets, only those in the Theme and Blend conditions made significant gains. Both higher and lower achievers in the Blend condition made especially strong gains.

During year two we also developed performance-based alternative assessments relevant to the content of the Quincy Market unit. Our intent was to identify benefits of instruction that would not be identified by traditional multiple choice achievement measures. Pre- and posttest scores for these problem-solving situations indicated that the Thematic and Traditional treatment groups did not differ significantly. However, both the Thematic and Traditional groups scored significantly higher and made steeper gains

than the Blend group.

Two master's thesis projects conducted under the auspices of the project were conducted during year 2. These activities were designed to assess the effects of alternative assessment on learning outcomes, and to determine whether or not the Thematic treatment influenced the use of mathematics communications. The performance-based assessment of learning in the domains covered by the Quincy Market curriculum was a product of one of those theses (O'Donnell, 1995). O'Donnell arranged for some Thematic classes to receive repeated exposure to alternative assessments during instruction on the Quincy Market theme. Other Thematic classes received conventional assessments. O'Donnell found that students who received repeated exposure to performance-based, alternative assessment did as well on a conventional multiple-choice posttest as did students who were not exposed to the alternative assessment practice. This was an expected outcome. But contrary to O'Donnell's hypothesis, performance of students in the alternative assessment condition on a performance-based posttest was not superior to that of students who did not receive repeated exposure to the alternative assessment practice. However, the results may be suspect because of multiple disruptions on the day of posttesting. (A substitute teacher who was unable to maintain order was in charge of one of the classes in the alternative assessment condition at the time of posttesting. The testing situation was further disrupted by the presence of guests who visited the class during testing.)

The form of alternative assessment used by O'Donnell requires students to communicate their thinking and problem-solving processes. There has been speculation and

some evidence that this form of assessment may disadvantage minority students whose native language is not the language of instruction to an even greater extent than conventional assessment. O'Donnell did not find this to be the case.

The second thesis study (Moss, 1995) investigated the proposition that assessment practices will have an effect on student's mathematical writing. He studied the effects of conventional versus alternative assessment practices on student's mathematical writing using the variables of Vocabulary-Count, Total-Word-Count, and (Mathematics) Vocabulary-Word-to-Total-Word-Ratio. He found no support for his hypothesis, although the trend was in the hypothesized direction and the somewhat low statistical power of his analysis (a 30 percent chance of missing a significant result at alpha level .05) leaves open the possibility that he was unable to detect a real result.

Quincy Market data for 1995-96 were analyzed by means of a 2 (sex) x 3 (treatment) analysis of covariance (ANCOVA), with CTBS Language Total and Mathematics Concepts and Applications as the covariates. The Trials by Treatment interaction was significant, F (2, 289) 4.327, $p < .01$. There were no other significant effects. As shown in Figure 7, the trials by treatment interaction results from differential pre- to posttest gains by the different treatment groups, with the Blend group making the largest gains, followed by the Theme group. It should be noted that one Blend teacher who taught Thematic classes during the previous year used the Quincy Market unit during the year under consideration. The Blend results in the present analysis were substantially influenced by scores in her classes.

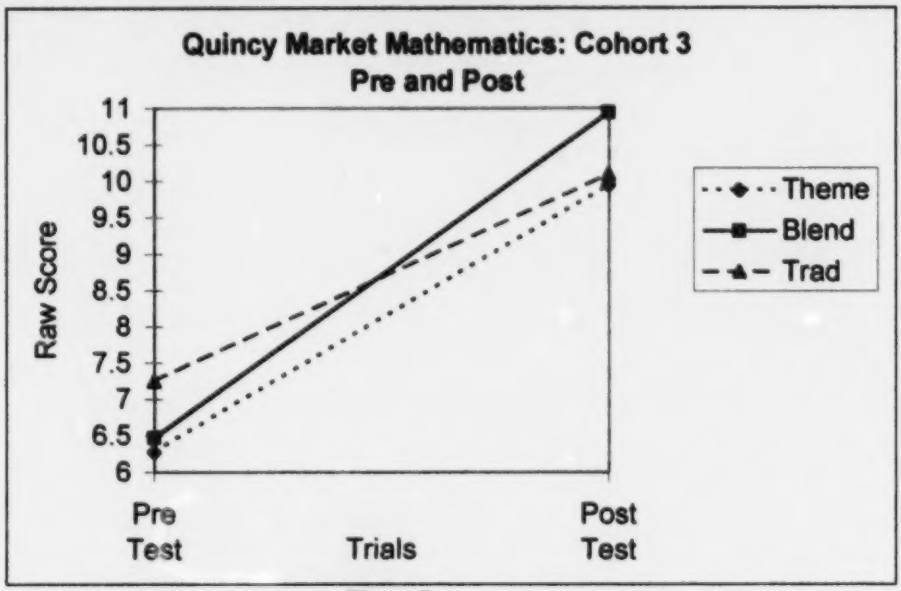

Figure 7

Grade Eight Achievement

The first cohort of students involved in the CAPP program at Washington Middle School made substantial gains against the national norms on the CTBS Mathematics Concepts and Applications test during the seventh grade. Thus, their grade 8 pretest scores (based on testing at the end of 7th grade in Spring 1994) scores, were quite high. Demonstrating additional gains against test norms during grade 8 would be a substantial challenge. Therefore, it was not surprising that the analysis of data for cohort one's 8th grade experience showed declines in NCE scores for the Traditional and Blend groups. The Thematic treatment group, however, maintained its position in relationship to national norms from pre-to posttest. Further analyses showed less positive effects for Latino students, compared with Euro-American students, in all three treatment conditions.

As was the case for cohort 1, Spring (i.e., June, 1995) scores were employed as the pretest for 8th graders of Cohort 2 to assess achievement results from the final year of

the project (1995-96) and, as before, the large gains made during the previous year made it a challenge to achieve similar gains during grade 8. Neither the Blend nor the Traditional treatments overcame this challenge. A 2 (trials) by 3 (treatments) by 2 (sex) ANOVA showed significant main effects for treatments, F (2, 236) 18.01, $p < .001$ and for trials F (1,236) 18.27, $p < .001$, as well as significant trials by treatment, F (2, 236) 7.29, $p < .001$, and trials by sex, F (1, 236) 8.58, $p < .005$, interactions. Figure 8 displays the trials by treatment interaction.

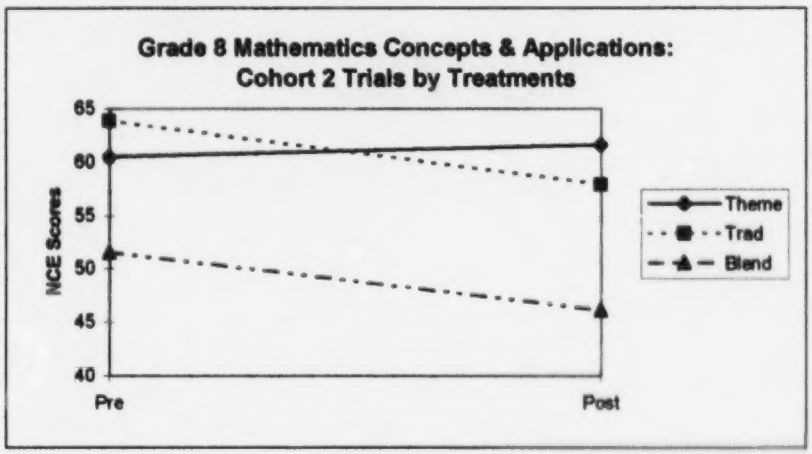

Figure 8

The significant Trials by Treatment interaction is accounted for by the fact that the Blend and Traditional treatments evidenced declines in Mathematics Concepts and Applications scores during the period from June 1995 to June 1996, whereas the Thematic group showed a slight increase.

As shown in Figure 9, the Sex by Trials interaction is attributable to rather steep pre to posttest declines for eighth-grade females, compared with only a slight [and statistically non-significant] decline for males.

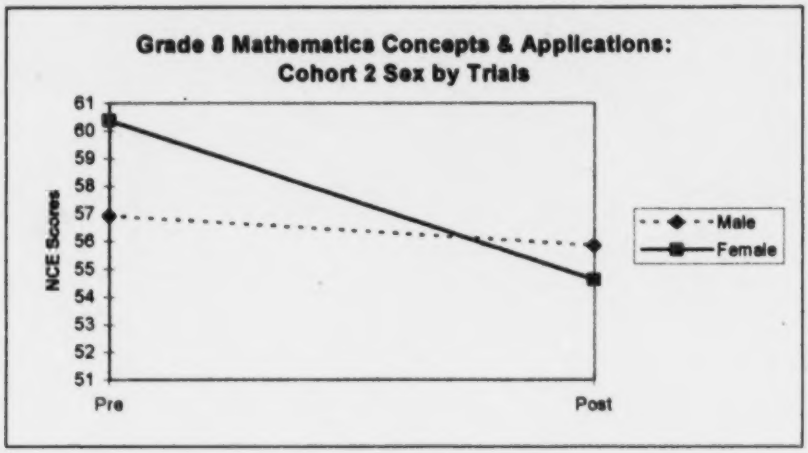

Figure 9

A 2 (trials) by 3 (treatments) by 2 (sex) ANCOVA, with CTBS Language Total as the covariate, was conducted to determine how mathematics achievement was affected by the treatments when English language achievement was held constant. In this analysis, the main effects for Treatments, F (2, 204) 4.51, $p < .01$, Sex, F (1, 204) 10.87, $p < .001$, and Trials, F (1, 204) 20.19, $p < .001$ were significant, as were the Trials by Treatment, F (2, 204) 7.24, $p < .001$, and the Trials by Sex, F (1, 104) 15.63, $p < .001$ interactions. Figure 10 depicts the treatment by trials interaction. The results reinforce those from the previous analysis; scores declined for the Traditional and Blend groups, whereas there as a slight increase for the Thematic group. The results also indicate that, even with statistical control for entry level language achievement, the scores of girls declined (relative to national norms) more sharply that did those of boys. The results are shown in Figure 11. Statistical control for entry level language achievement did not alter the finding of marked decline for eighth-grade females, relative to national norms and to the

achievement of their male peers.

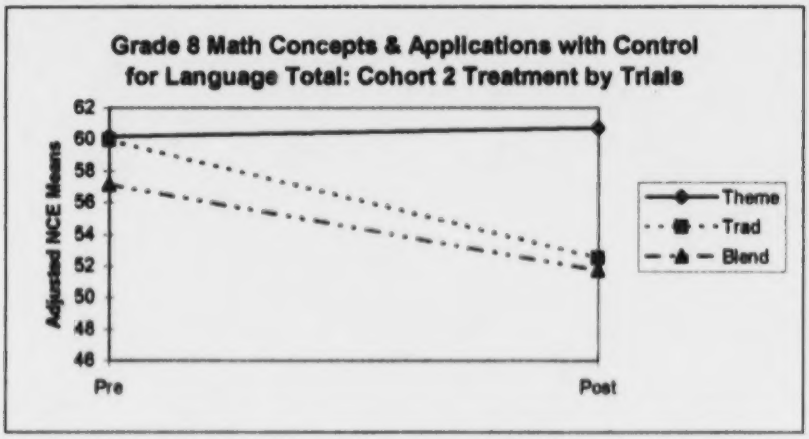

Figure 10

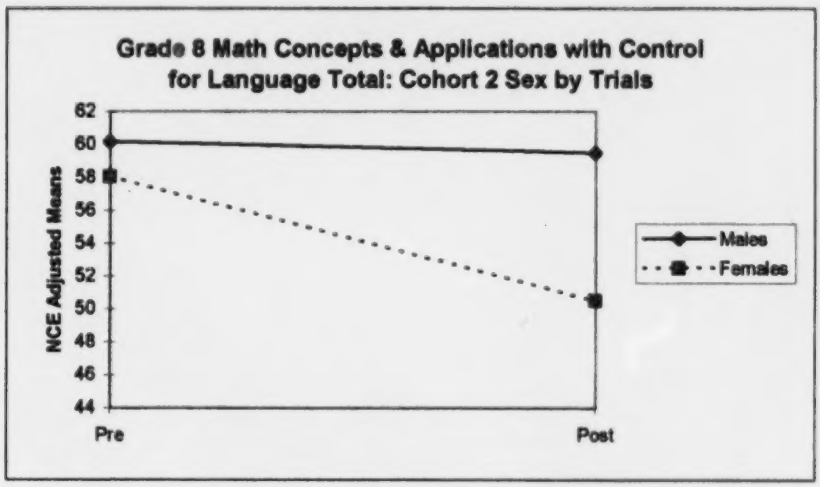

Figure 11

Mathematics in Manufacturing Curriculum: During the second operational year of the project, some eighth grade classes participated in a thematic curriculum, *Mathematics in Manufacturing (M in M)*, developed in a collaborative effort among the business partner, Radionisc, teachers at Washington Middle School, and mathematicians from the University partner in the project. Criterion referenced testing showed significant achievement gains for the classes

involved. The curriculum underwent continued development and refinement during the final operational year of the project.

The *Mathematics in Manufacturing* criterion test used in the final year of the project sampled the specific mathematics material that was covered in the *Mathematics in Manufacturing* unit. The material covered in the *M-in-M* unit were common to the regular 8th grade curriculum as well, although the mode of instruction differed. The test was administered to all 8th grade classes and the data were analyzed by means of a 2 (sex) x 3 (treatment) analysis of covariance (ANCOVA), with CTBS Language Total and Mathematics Concepts and Applications as the covariates. The main effect for condition, F (2, 325) 79.065, $p < .001$, and the trials by condition interaction, F (2, 325) 16.189, $p < .001$, were significant. The interaction is depicted in Figure 10. With statistical control for prior mathematics and language achievement, both the Blend and the Thematic treatment conditions demonstrated significant pre- to posttest gains, whereas the Traditional group did not. These results are shown in Figure 12.

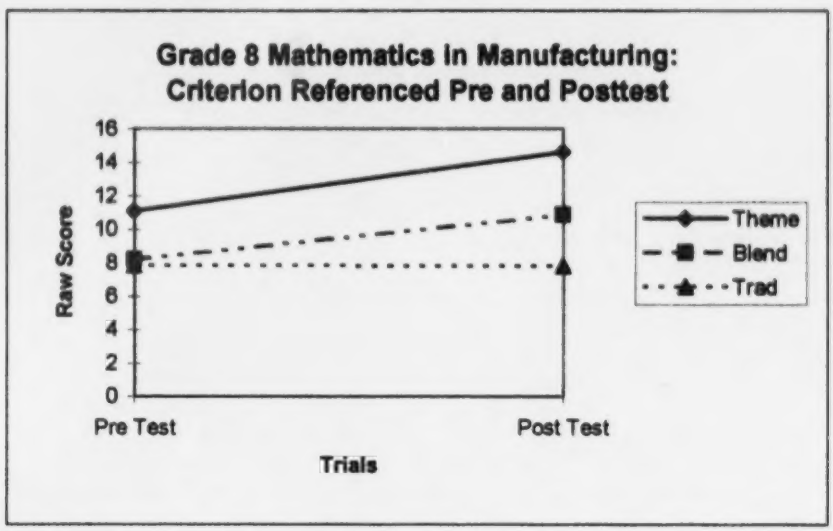

Figure 12

Mathematics Attitudes and Self-Perceptions

The several mathematics attitudes variables were examined separately for seventh and eighth grade by means of a series of 2 (trials) by 3 (treatment groups) by 2 (sex) ANOVAs followed by post hoc multiple comparisons.

Grade Seven Attitudes. The ANOVA on seventh graders' Mathematics Self-Concept (MSC) showed significant main effects for Treatment, F (2, 189) 3.27, $p < .05$, Sex, F (1, 189) 4.16, $p < .05$ and Trials, F (1, 189) 4.98, $p < .05$. The Treatment by Trials interaction was marginally significant, F (2, 189) 2.91, $p = .057$. Pre- and posttest means for the treatment groups are shown in Figure 13. The Traditional group began the year with significantly higher MSC scores than the other two groups, but the MSC scores of these students tended to decline over the course of the year, whereas those of students in the Blend and Thematic groups were stable. The sex effect was explained by overall higher scores for males (pretest mean = 19.02, posttest mean = 17.99) than for females (pretest mean = 17.52, posttest mean = 17.24).

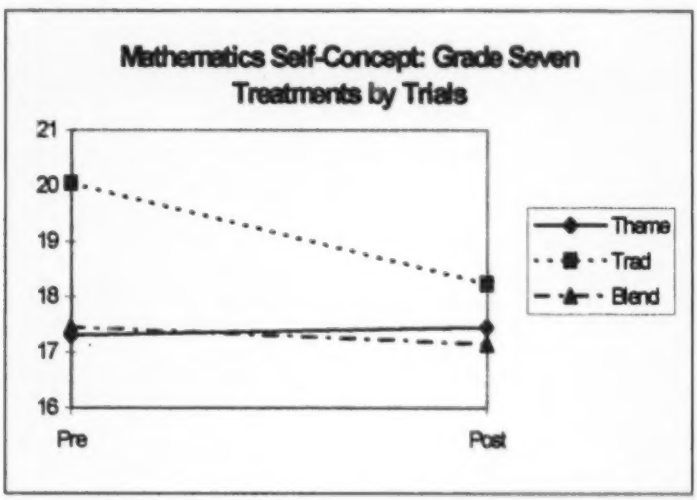

Figure 13

BEST COPY AVAILABLE

The analysis of *Mathematics Attitudes* (MATT) revealed a significant effect for Treatment. There were no other significant effects. Scores of students in the Thematic treatment begin lower and stayed lower than those of students in the other two treatment groups. The *Mathematics Attitudes* scores of Blend students were constant from pre- to posttest, and the declining trends indicated in Figure 14 for the Thematic and Traditional groups were not significant.

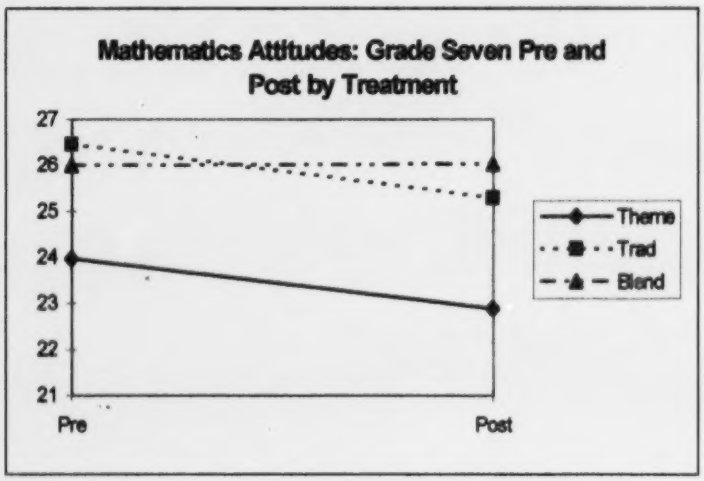

Figure 14

The only significant finding from the analysis of Parental Encouragement for mathematics learning was a main effect for trials, F (1, 148) 5.53, $p < .05$. Overall, student's ratings of the Parental Encouragement increased over the course of the school year.

Grade Seven Self-Perceptions. The repeated measures ANOVA for the Challenge-Seeking/Persistence variable yielded a significant main effect for Trials, F (1, 207) 5.29, $p < .05$ and a significant Treatment by Trials interaction, F (2, 207) 4.64, $p < .01$. As depicted in Figure 15, the Thematic and Traditional groups declined on this variable over the pre to posttest interval, whereas the Blend group was relatively stable.

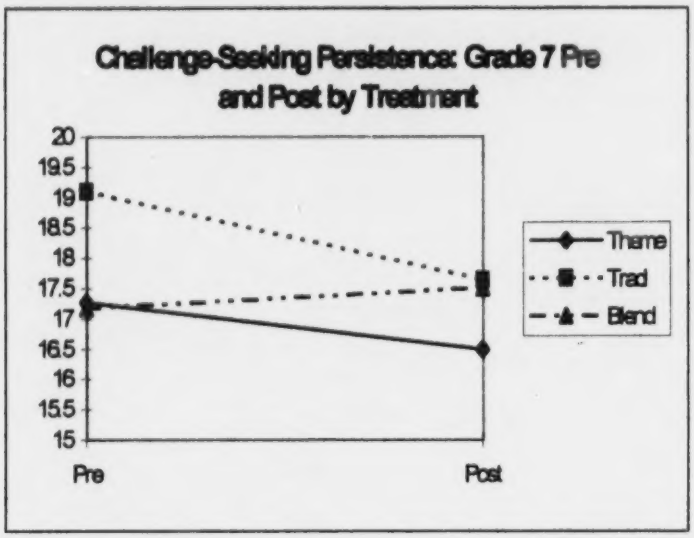

Figure 15

The only significant effect found in the analysis of the Scholastic Competence (SC) variable was three-way interaction for Trials by Treatment by Sex. Scores of males in the Thematic group increased whereas their female peers evidenced a decline in their self-perceived scholastic competence. Traditional Males also increased in their self-perceived Scholastic Competence, whereas females in that treatment group showed only a slight increase. The pattern was reversed in the Blend group, where females showed gains while the males showed no change between pre- and posttesting. The results are shown in Figure 16.

The analysis of seventh graders' self-perceptions of Social Acceptance yielded significant main effects for Treatments, F (2, 178) 4.72, $p < .01$, and for Trials, F (1, 178) 5.43, $p < .05$. No other effects were significant. Overall, Social Acceptance scores for Blend students were lower than those of their peers in the other two groups. Scores of the Blend (pre = 17.34, post = 18.03) and Traditional (pre = 18, post = 19.04) treatment groups increased over the pre- to

posttest interval, whereas scores of the Thematic group did not (pre = 19.02, post = 19.00), but the treatment by trial interaction was not significant.

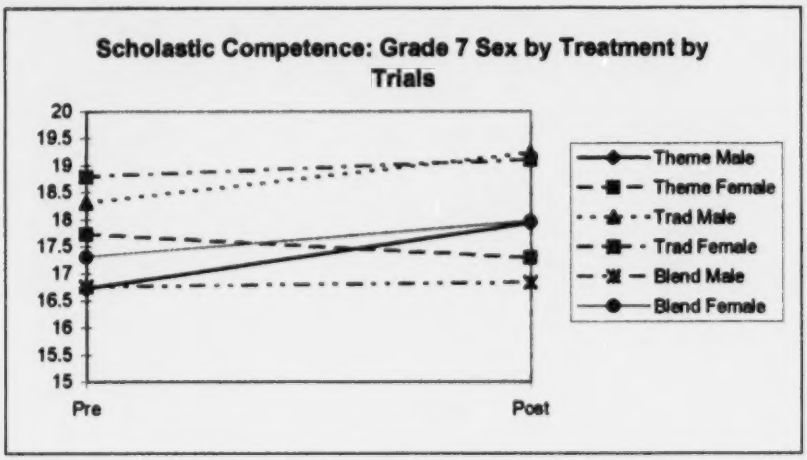

Figure 16

The analysis of Global Self-Worth self-perceptions yielded a significant main effect for Treatment, F (2, 201) 7.39, $p < .001$, and a significant Treatment by Trials interaction, F (2, 201) 3.07, $p < .05$. Figure 17 shows the interaction, which is explained by a pre- to pottest decline for

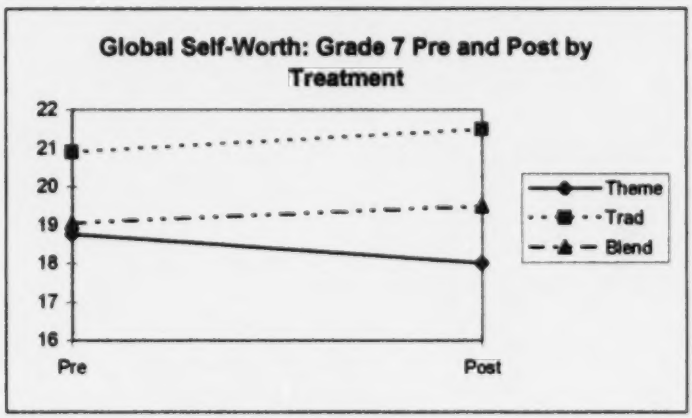

Figure 17

the Thematic treatment group, contrasted with slight increases for the Traditional and Blend groups.

Grade Eight Attitudes.

The only significant finding from the analysis of Mathematics self-confidence for eighth graders was a main effect for sex, F (1, 241) 4.59, $p < .05$. Females expressed less self confidence in their mathematics ability at both the beginning and end of the school year than did their male peers.

The analysis of Mathematics Attitudes also indicated a significant main effect for Sex, F (1, 198) 5.39, $p < .05$, as well as a significant main effect for Treatments, F (2, 198) 3.10, $p < .05$. There were no other significant effects. Females expressed less favorable attitudes toward mathematics than did males at both the beginning and the end of the school year. The most positive Mathematics Attitudes were expressed by students in the Blend treatment group (pre = 23/59, post = 24.17), followed by those in the Thematic group (pre = 22.85, post = 22.68), and then by those in the Traditional group (pre = 21.72, post = 21.56). As indicated by the lack of a significant Trials effect, Mathematics attitudes did not change over the course of the school year in any of the treatment groups.

The analysis of Parental Encouragement for mathematics learning yielded a significant three-way interaction for Treatment by Trials by Sex, F (2, 202) 4.24, $p < .05$. The interaction is depicted in Figure 18. Females who were in the Traditional treatment group showed increases in their perceptions of the encouragement for learning mathematics they received from their parents, whereas there were not changes for females in the Blend and Thematic groups. In contrast, males in the Traditional treatment group showed declines on this variable, whereas males in the Blend and Thematic groups showed increases.

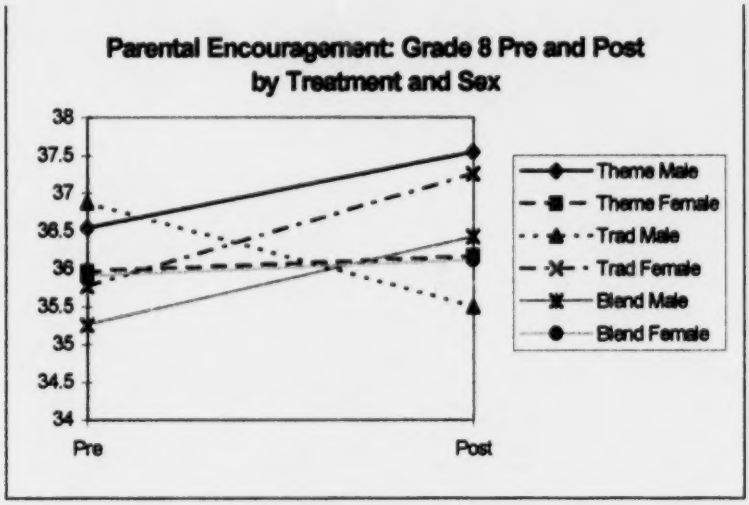

Figure 18

Grade Eight Self-Perceptions.

The analysis of the Challenge-Seeking/Persistence variable yielded a significant main effect for Sex, F (1, 296) 4.73, $p < .05$. In addition, the Sex by Trials, F (1, 296) 3.83, $p < .05$ and Treatments by Trials by Sex, F (2, 296) 3.45, $p < .05$ interactions were significant. The interactions are shown in Figure 19.

Males in the Thematic treatment scored higher than their female peers, neither of whose mean scores changed over the course of the year. In the Traditional treatment group, females began the year with higher scores than males, but whereas scores of males increased, those of females decreased. Therefore, at the end of the year the mean for females' scores on this variable were higher than those of their male peers. In the Blend group, males scored higher on *Challenge-Seeking/Persistence* than females on both the pre and posttest. Means for both males and females remained stable from pre to posttesting.

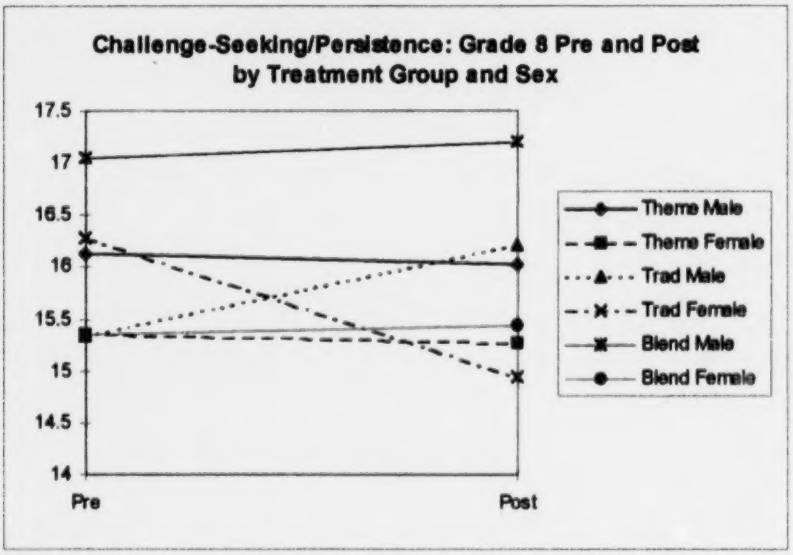

Figure 19

The only significant finding from the analysis of self-perceived Scholastic Competence was a significant Sex by Treatment interaction, F (2, 259) 2.99, $p < .05$. As can be seen in Figure 20, overall SC scores of Blend females was lower than those of their male counterparts.

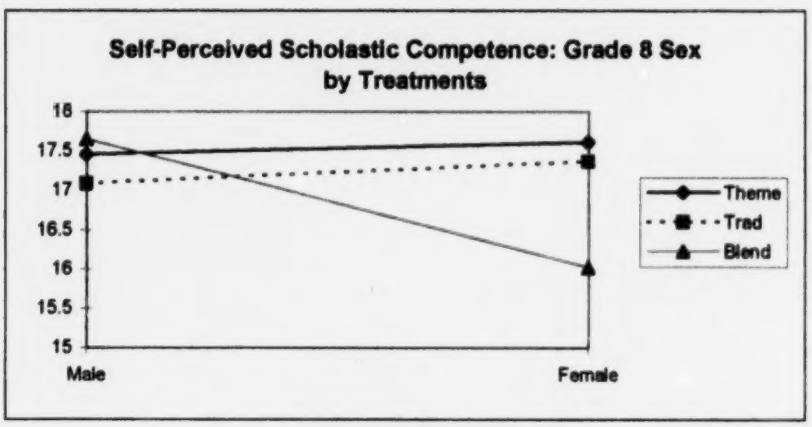

Figure 20

whereas females' had slightly higher scores than their male peers in both the Thematic and Traditional treatment groups.

The analysis of Self-perceived Social Acceptance identified a significant Sex by Trials interaction, $F(2, 248)$ 7.53, $p < .001$. The interaction, contrasting a pre to posttest increase for females with relatively stable pre to post scores for males is displayed in Figure 21. The analysis of Global Self-worth for 8th graders found no significant effects.

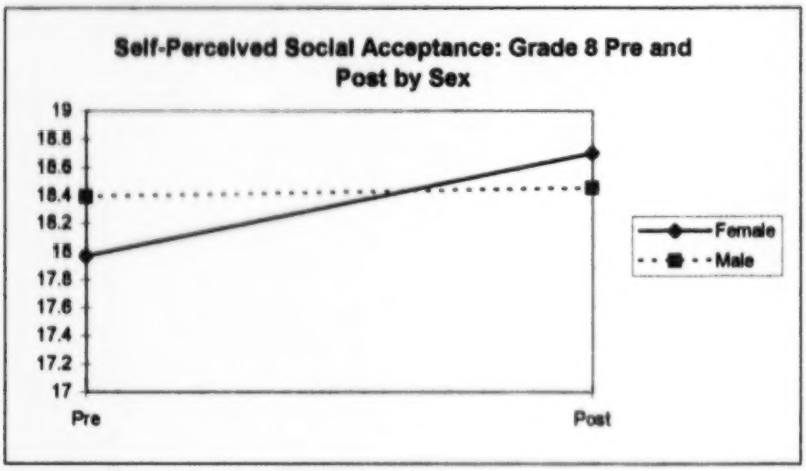

Figure 21

Student Interactions in Small Instructional Groups

An in-depth study was carried out in three, eighth-grade thematic classes. This dissertation examined how small group interactions and individual background characteristics of middle school students relate to participation in mathematics activities (St. John, 1996). The central questions were: what individual background characteristics (gender, ethnicity, prior mathematics achievement, prior language achievement, and peer status) are associated with student participation in mathematics activities, what role does mathematical discourse play in group interactions, and do other group members influence a student's behavior in the group? Students

(n=72) from three thematic eighth-grade classes were videotaped as they worked in small groups (four students per group) on mathematics assignments in their classroom. The tapes were coded for students' participation, the complexity and quantity of mathematics discourse, and exclusionary behavior observed during the activity.

Although girls spoke more about mathematics than boys, both boys and girls spoke and participated at equally complex levels. Participation and mathematics talk were significantly related to the participation and mathematics talk of other group members for girls, but not for boys. Exclusionary comments were directed to boys by other members in their group more often than such behaviors were directed toward girls. When boys were thus excluded by their group members from the assigned activity, they participated less. As with girls, the participation and mathematics discourse of Latino students was significantly related to the activity of other group members, whereas there was no relationship for Euro-American students. The best predictor of participation and mathematics discourse was the participation and discourse of other members in the group, not background variables such as prior mathematics achievement or motivational variables such as goal orientation.

These findings provide evidence that the participation of most students in group mathematics activities is substantially influenced by other group members. Although the personal cultural and social histories of students were related to how they interacted in groups in mathematics activities, the behavior of other group members played the greatest role in how individual students participated in mathematics activities. Framed within a sociocultural perspective, these findings support the notion that motivation, when viewed as participation, is assumed to be constructed in activity.

This study has important implications for teachers implementing group work in their classrooms. First, students with limited mathematics and language achievement are able to engage in mathematics work in small groups. Second, given the importance of other group members to the participation and mathematics discourse of girls and Latino students, it is important for teachers to monitor group interactions. Likewise, for boys and for Euro-American students, negative comments about their contributions negatively impacts their participation in the activity. This study suggests that the participation of the group members plays a key role in how individuals respond. Hence it is critical that teachers make efforts to ensure that the group is engaged in the assigned activity. Training students how to work with each other may assist with the problem of boys and Euro-American students participating less when they experience being excluded from the activity.

Predicting Mathematics Achievement

Analyses of attitudinal and self-perception variables indicated that the best combination of predictors was *Mathematics Self-confidence*, *Mathematics Attitudes*, and *Perceptions of Parental Support for Mathematics Learning*. This pattern held true at both the seventh and eighth levels. For seventh grade, pretest scores on these variables accounted for nearly 17 percent of the variance in student performance on the CTBS Mathematics Concepts and Applications posttest, $R\,(3,215) = .408$, $R^2 = 16.6$, $p < .001$. Eighth graders' pretest scores on these variables accounted for 15 percent of the variance in CTBS Mathematics Concepts and Applications posttest scores, $R\,(3,181) = .387$, $R^2 = .15$, $p < .001$.

To put these findings in perspective, recall that there was no clear pattern of

change for the different treatment groups over the course of the school year. Whereas we could not identify interpretable differences in the effects of different instructional treatments on attitudinal or self-perception measures, the importance of attitudes and student perceptions was born out by the ability of a subset of these variables to explain a substantial proportion of the variance in mathematics achievement.

Qualitative Outcomes of CAPP Project

Although the evaluation effort focuses most directly on student outcomes related to the learning of mathematics, it is also important to examine how the presence of the project and the activities it has brought to Washington Middle School has influenced the institution. In fact, these effects may be the more important outcomes to document at this point in the life of the project, because it is a well established generalization among educators who have been involved in efforts to effect curricular and instructional change that it takes from three to five years to institute even the most basic reforms. It is therefore crucial to understand what changes are underway that could eventuate in improved student learning.

The CAPP project focused its time and resources on assisting thematic teachers in developing and implementing thematic units. This included collaboration between the university team and teachers in designing group work, encouraging inquiry-based instruction, revising and supplementing existing thematic units, and creating the *Mathematics in Manufacturing* curriculum at the eighth grade level. The Business Partner, Radionics, Inc., was (and continues to be) a vital contributor to the *Mathematics in Manufacturing* development effort. Except for occasional classroom observations and infor-

mal conversations, being represented on the CAPP Project Steering Committee, and participating in a number of in-service opportunities provided under the auspices of the CAPP project, teachers in the blend and traditional curriculum groups received scant attention. Even so, we believe there is clear evidence of a contagion effect that has affected teachers in all three treatment groups. For example, the project emphasizes practices that engage heterogeneous groups of students in instructional conversations. Teachers from all treatment groups attended a CAPP Project sponsored day of in-service on the topic of *instructional conversations* (see Tharp & Gallimore, 1988). Several teachers representing each of the treatment groups sought funds to attend in-depth training in *Complex instruction* to extend the topics that were opened by the CAPP sponsored in-service. While all teachers appeared to benefit from this training, the effect was most evident in the teaching style of a traditional teacher. Based on yearly classroom observations (conducted by the same research assistant), we noted that the traditional teacher had broadened his approach of traditional question, answer, and evaluation of students, to include the use of inquiry-based questioning. He also began using demonstrations to introduce topics, such as using an actual scale, yardstick, and meterstick to introduce the topic of measurement.

It should also be noted that during the final year of the project, a teacher who had taught in the Thematic approach during the first two years taught Blend classes, bringing to her Blend instruction the insights acquired through her Thematic teaching experience. Additionally, this teacher utilized some of the thematic curricula during the year.

Another example of contagion comes from one of the blend treatment teachers.

During the 1994-95 school year, she dramatically increased the use of Renaissance units so that they now comprise the bulk of her curriculum. Some of these units, such as one on Baseball, spanned several weeks. The blend teacher introduced it to the thematic teacher who adopted its use. This increased use of thematic-like units blurs the differentiating features of the thematic and blend groups. Among the thematic teachers, one was initially reluctant to participate in curriculum supplementation and development during the first year of implementation. However, during the spring of the second year, she informed the research team that she had developed an economics unit and was utilizing it with her students. Furthermore, the teacher designed some of the assignments to ensure students would link the economics unit to the subsequent Quincy Market unit. The unit required students to take their mathematics learning out into the community by interacting with local businesses. Each of these teachers expanded their teaching in some way that was not a required nor expected part of their participation in the research project. Additionally, these explorations seemed to draw upon the atmosphere of inquiry set by the research project.

In addition to the contagion effect, we have been struck by the increase in communication among teachers within the school and district at large. Within the school, the project has encouraged rather than insisted upon communication among teachers. Initially, teachers began to informally discuss curricular materials (i.e., sharing of Renaissance units between a blend and thematic teacher). In the second year the mathematics teachers embarked on a more formal communications procedure by scheduling peer observations. These observations began among three teachers involved in teaching the-

matic curriculum. One teacher began to use thematic units with transitional Spanish speaking students during the second year. Although arrangements were difficult, she began sharing information and observing both seventh and eighth-grade thematic teachers and vice versa during the final year of the project.

The CAPP project has also stimulated discussion about curriculum. During the 1994-95 academic year, after much discussion, the teachers created a scope and sequence of concepts and skills for the middle school level. They took this to the other middle schools, elementary schools, and high schools in an effort to stimulate the development of a comprehensive, well articulated, K-12 district-wide mathematics program. During the spring, they invited high school mathematics teachers to tour the Radionics manufacturing plant and learn about what the thematic approach to teaching mathematics entails. Previously unwilling to make significant changes to the high school sequence, the overture made by the middle school teachers generated a great deal of interest by high school teachers in the thematic approach. These teachers have been negotiating changes to the high school courses and offerings to bring them more in line with the type of mathematics learning many students experience at Washington Middle School.

It seems evident that these changes in communication and engagement in the design and specification of curriculum were encouraged and supported either directly or indirectly by the presence of the CAPP project. Although they are not directly measured, the increase in communication and exploration of teaching styles and curriculum are essential for educational reform to be sustained.

Employees of Radionics, our CAPP business partner, have also benefited from the

CAPP project in unexpected ways. Student visits to the manufacturing plant are workdays for the employees and students. Each student works with employees in two different areas of the plant during the course of the day. The employees clearly enjoy the opportunity to pass along their knowledge and skills to students. After the first few student workdays, several employees reported a new interest in identifying and verbalizing how mathematics is used in their work. After working with students and encouraging them to pursue their education, a few employees reported they were inspired to take coursework themselves and apply for more advanced positions within the company.

Along with employees, students appreciated the thematic approach to teaching mathematics. One aspect of the this approach includes an emphasis on group problem solving. During interviews with eighth-grade thematic students, one student commented, "We've done more group work in a month this year than I have in a whole year in another class." This group emphasis was supported by most students reporting their preference for group work over individual work in mathematics. Regardless of ethnicity or academic achievement, students commonly said things such as, "I learn more in a group than working by myself because you might be in rows doing individual work and the teacher might say something and you don't understand as opposed to being in a group, then the people in your group can help you." Another student said that one of the benefits of group work is the "We don't die out while we're doing our work and fall asleep or something." Although the students were not directly queried about thematic instruction, several students volunteered that they preferred thematic work to "...straight numbers and calculations because it's easier, and the creativity of it makes me think more."

There is evidence that the reforms the CAPP project supported are being sustained. In the area of curriculum development and community involvement, nearly all eighth grade students worked at the Radionics plant during the 1996-1997 school year. Teachers decided to standardize the scope and sequence across grades. Their rationale included accepting greater accountability for student outcomes, making it easier for students to move from one class to another, and making it easier for teachers to share ideas. One teacher said that she had rearranged the sequence of presenting topics because she agreed with the importance of covering material at similar times to her colleagues. She said that it has facilitated sharing of assignments among teachers.

Conclusions and Discussion

The use of standardized tests is often cited as a barrier to instructional change. Teachers who are encouraged to adopt instructional reform in mathematics often report reluctance to depart from the traditional curriculum because they fear that their students will then not perform well on the standardized measures on which their own effectiveness is often judged. In the present evaluation we used multiple measures, including alternative assessments designed to determine if Thematic instruction delivered educational value beyond that measured by standardized tests, and we employed measures specifically constructed to assess the skills and concepts being taught in the alternative curricula. But we also used standardized tests. Since these measures played a dominant role in the evaluation, a word of explanation is in order.

Most teachers who participated in teaching the alternative curriculum did not subscribe to the view that their students' performance on traditional standardized tests would

suffer as a result of reform-oriented approaches to mathematics instruction. However, one teacher who taught 7th grade theme classes throughout the duration of the project did not subscribe to the idea that reform-oriented instruction would not be less effective for her students than traditional practices. Moreover, all of the teachers were skeptical about the pedagogical wisdom of heterogeneous grouping in mathematics classes (whatever the curriculum), they were otherwise convinced that, if the program were effective, students would also do as well as students taught by traditional methods in learning the skills and concepts that are sampled by standardized achievement tests.

The overall results of the evaluation seem to support the teachers' confidence. The use of a thematic curriculum did not result in poorer achievement on standardized tests than would have been the case if those same students had experienced a more traditional curriculum. This was true for seventh graders during all three years for which data were collected.

The data suggest that the project school is doing an especially effective job at the seventh grade level. Results from the first operational year of the project showed positive gains across treatment conditions. The data examined for the second year were even more impressive, and the pattern of increases against national norms was maintained during year three. Data from the fall to June testing period for the final year of the project showed parallel gains for the Thematic and Blend treatments, both of which were greater than the gains registered by the students in the Traditional treatment. The improvement in rate of change for the Blend group during year 3 was particularly striking. A class-by-class inspection of the data suggested that the improvement was due in large

measure to achievement by students in Blend classes that were taught by a teacher who had taught the Thematic approach during the previous two years. As a Blend teacher, she continued to use the seventh-grade Quincy Market materials, with extensive elaboration and language-rich additions of her own design. The results of these activities were reflected in achievement based on the criterion-referenced test for Quincy Market, on which the Blend as well as the Thematic group demonstrated greater progress than did the Traditional treatment group. Features which the seventh grade versions of the Thematic and Blend approaches shared in common, as they were actually implemented in year three, included careful attention to scope and sequence (thus ensuring students' opportunity to learn the full range of content for their grade), and language enrichment. We speculate that these features helped to account for the excellent achievement results of these two treatments.

One of the most dramatic findings was one we did not set out to look for. A subset of the 3rd cohort of 7th grade classes experienced three different teachers during the course of the year. The initial teacher was a regular member of the mathematics faculty, but classes were reconstituted early during the year to accommodate actual enrollments. A teacher with an emergency credential was then assigned to teach these three classes. The regular mathematics faculty considered him to be poorly qualified for the assignment, but they tried to assist him and he was receptive to their help. But just as he was beginning to feel more comfortable with the assignment, a third teacher, recruited in Mexico, was assigned to this subset of classes. Members of the mathematics faculty considered him unqualified, and there were many complaints from students and parents as

well regarding his competency, or lack of it. Some said he was unresponsive to their offers of assistance. He had little control over his classes but remained in the assignment, nevertheless, for the remainder of the school year.

We recite this history as a context for the achievement results we observed. Although these classes were nominally assigned to one or another of the treatment conditions, on-site observations made it clear that what transpired in these classes could not be legitimately classified as representing any of the planned treatments. These classes were therefore treated in the analyses as a separate treatment. The analyses showed sharp declines in achievement over the course of 7th grade for students in these classes.

When we met with the teachers at the end of the year to share the evaluation results and to hear their interpretations, some teachers argued that a similar phenomenon could be observed in almost any school in the state. They attributed the sad achievement plight of students in these classes to the fact, or what they took to be a fact, that school districts typically take a very conservative approach to hiring, preferring to err on the side of having too few teachers assigned to a school rather than too many. When actual enrollments are known the administration then scurries to find additional teachers, but by that time the best qualified teachers have already found positions. If the viewpoint represented at the meeting is accurate, then countless numbers of students in California may be shortchanged in their educational experience. School administrators should think seriously about hiring practices and their consequences.

The 8th grade data for cohorts 1 and 2 (program years 2 and 3) were less encouraging than those for the 7th grade. Both years showed a pattern of decline for eighth

grade students, with the exception of those in the Thematic treatment. These declines in relation to national norms during eighth grade remain to be explained. Clearly, the very large gains against national norms during grade seven would be difficult to maintain, but even here the results for the thematic treatment are encouraging in that the pretest to posttest analysis demonstrated that in year two the Theme group held their own against the norms, whereas the other two groups declined. In year three the Theme group made small gains against national norms, whereas the other two treatments displayed declines.

When we met with the mathematics faculty at the end of both years 2 and 3 to discuss the data and to seek interpretations from the teachers' perspective, a variety of possibilities were suggested. One teacher expressed the belief that the results for the Blend approach might be explained by the lack of attention to the scope and sequence of mathematics material covered in the program. There was some support for this idea among other teachers, but the one teacher who was a strong advocate for the approach being used in the 8^{th} grade Blend was absent from this part of the discussion. The 8^{th} grade Blend approach relied heavily on the use of *Math Renaissance* materials which emphasize hands-on activities for pairs of students or small groups. Based on informal classroom observations, we were of the opinion that this teacher did an excellent job of implementing the Renaissance approach, but it does seem plausible that the activities around which the materials are organized may provide incomplete coverage and/or linkage of the concepts and skills students are expected to study in the eighth grade. If this is true, students who receive this kind of instruction might have less opportunity than their peers who experience instruction that was explicitly attentive to scope and sequence to

learn the expected content. Given the popularity of Renaissance material as a way of addressing some of the recommendations of the mathematics reform movement, the "scope and sequence hypothesis" warrants closer examination.

Students in the Traditional treatment had the benefit of instructional coherence based upon the scope and sequence embodied in a textbook, but they learned mathematics in a decontextualized manner that tends to treat concepts and skills as discrete units with little real world application. The *Concepts and Applications* test used as the achievement criterion does require that students make some application of concepts and skills, perhaps accounting for the achievement decline among 8th graders in this treatment group. However, this explanation would not be consistent with the fact that 7th graders in the Traditional treatment demonstrated healthy achievement gains on this same criterion measure.

We were especially curious about the rather steep decline in mathematics achievement among 8th grade girls. When we discussed the evaluation results with the teachers they expressed disappointment with this pattern, but some said they were not surprised. They speculated that changes in peer pressure and sex role expectations began to exert a strong effect in 8th grade. Interest in boys became especially strong among the 8th grade girls, who began to regard it as inappropriate for girls to be good students, especially in mathematics. Their level of effort fell off accordingly. Other teachers seemed to agree that this was a widespread pattern in the school. The pattern was especially discouraging for teachers who have given significant attention to issues of gender equality in mathematics and who have gone to special efforts to provide models of women in

mathematics and in mathematics-based professions.

As anticipated, 8th grade students in the Thematic treatment achieved especially well in relation to the Traditional treatment group on the Mathematics in Manufacturing test. Although the test included skills and concepts that are common to the 8th grade curriculum, the problems were highly contextualized and required fairly complex problem-solving processes. The Blend group, which also received contextualized instruction, did well on this criterion test as well, suggesting that contextualization, rather than the specific content, contributed to successful outcomes.

Experiments with performance-based alternative assessment produced some interesting findings. Students who were exposed regularly to performance-based tests performed as well on conventional multiple choice tests as did students who were not exposed to alternative assessment. But neither did they demonstrate superior performance on alternative assessments, compared with students who were not tested with alternative assessment instruments on a regular basis. These students did not appear to be disadvantaged by their lack of practice with performance-based alternative assessment approaches. From these results it appears that a practice effect from experience with this particular form of assessment was not a dominant factor in student test performance. These experiments also showed that non-native speakers of English were not disadvantaged, in relation to their native English speaking peers, on measures that required written communication of students' problem-solving thought processes.

Attitudes and self-perceptions seem to be especially resistant to change. The hypothesis that students would demonstrate improvements in mathematics-related attitudes

and self-perceptions as a result of participation in contextualized instruction that stressed hands-on activities and small group learning was not supported. Pre to post test results showed no clear pattern of advantage for one instructional approach over another. In fact, there were more declines than increases on these measures, perhaps reflecting the general nature of attitudes toward school among students who are entering adolescence. Importantly, although discouraging, we did find that girls, compared to boys, showed a decline in mathematics self-confidence. This held true for both 7th and 8th graders.

Although none of the instructional approaches appeared to have a generally positive influence on mathematics-related attitudes and self-perceptions, these constructs nevertheless appear to play a significant role in mathematics achievement. At both the 7th and 8th grade levels, Mathematics Self-confidence, Mathematics Attitudes, and Perceptions of Parental Support for Mathematics Learning functioned as significant predictors of mathematics achievement. The question of what it takes to achieve improvements in attitudes and self-perceptions at this level of schooling remains to be answered.

A major finding from intensive study of student interactions in small instructional groupings of 8th graders in the Thematic treatment showed students' level of content-relevant participation in small group learning activities is functionally related to the participation level of other students in the group. We would conjecture that students who have the opportunity to work in well-functioning groups, characterized by high levels of content-relevant participation, would eventually appropriate high levels of task-oriented mathematics problem-solving motives into their own motivational systems. The present study was much too brief to examine this speculation, but it should be a fruitful topic for further investigation. What does seem clear from

these findings is that there is wide variability in levels of student participation in small learning groups. Since content-relevant participation begets content-relevant participation, students should be provided with the training, monitoring, and coaching necessary to help them function as contributing members of learning groups.

Finally, it is important to note that the school's relationship with our business partner has continued beyond the life of the CAPP project. Although there has been a change in ownership of the business and some manufacturing functions have been transferred to overseas operations, students and teachers are still involved with the company. If anything, the numbers of students and classrooms involved has increased. In addition, there is more communication and exchange of ideas about curriculum and teaching than was the case when the project began. The teachers plan to make some changes in curriculum and classroom organization, on the basis of their expe-rience during the term of the CAPP project, but the essential commitment to an enriched cur-riculum remains.

Postscript

The mathematics teachers at the participating middle school read a draft of the final in-ternal evaluation report and met to discussion their reactions. In this section we report their ob-servations and additional analyses that were conducted to address issues that surfaced during their discussions.

The Issue of heterogeneous grouping continues to be a hot discussion topic at the school, with seven of the eight mathematics teachers indicating that they do not believe in the purported benefits of this grouping practice. Even though the evaluation data suggest that students in the treatments that involved heterogeneous grouping made achievement gains were equivalent to or

better than those attained by their peers in homogeneously grouped Traditional classes, the teachers are not at all convinced of the efficacy of heterogeneous grouping. Among at least some of them there is considerable resentment about the gifted (GATE) program, which they feel takes the "cream of the crop," leaving too few high achieving students in classes designated as "heterogeneous" to provide the student leadership and modeling needed for successful mixed ability small group work.

Earlier in our discussion we mentioned a discussion in which we were searching for an explanation of sharp declines in mathematics achievement among 8th grade girls. We reported on speculation that those declines might result from changes in peer pressure and sex role expectations the affect 8th grade girls. When they met to consider the draft of the final evaluation report, the teachers decided that we should cut that section. Their rationale was that the programs offered by the school district should hold students accountable and enable them to fulfill their possibilities in all areas of the curriculum. Given this philosophy, they felt that their earlier speculation seemed like an excuse. We have chosen not to eliminate this the original speculation because it provides an important hypothesis to be pursued in future research, especially if this pattern is found to be widespread among certain populations of female students. But we do register here the teachers' view that schools should be responsible for the achievement of their students, and that there should be no excuses for poor performance.

The major independent variables of interest in this evaluation were instructional practices/curriculum and grouping practice. An ideal evaluation design would have counter balanced these variables, so that each curricular approach was taught in both homogeneous and heterogeneous conditions. It was not possible to achieve such a design within the circumstances of the

school. Therefore, the effects of grouping cannot be separated from the effects of different curricula or instructional approaches. Thus, we have no way to determine whether the observed outcomes resulted from the curriculum, the grouping practice, or a combination of both. However, we are quite confident that the achievement gains made by students in the heterogeneously grouped classes were parallel to those of their peers in homogeneous classes. Consequently, it does not appear to us that their learning was hampered by heterogeneous grouping, and the achievement of 8th graders in the Thematic (heterogeneously grouped) treatment was superior to that of students in either the Blend (heterogeneously grouped) or the Traditional (Homogeneously grouped) treatment.

At least one teacher also wondered if students in the standardized test had insufficient ceiling to reflect the true extent of student learning among 7th graders in the Traditional treatment. Recall that the pretest mean for students in that treatment was significantly higher than pretest means for students in the other two treatments. If students in the top quartile were "topping out," so that the CTBS test was incapable of reflecting additional growth, the pre- to posttest gains for students in Traditional treatment would have been artificially attenuated. To test for this possibility, we conducted additional analyses of the 1995-96 (3rd Cohort) in which we did not include any student with a pretest score in the top quartile. The results of the 3 (treatments) by 2 (trials) repeated measures analysis of variance showed significant effects for treatment, F (2, 292) 6.43, $p < .01$, and for trials, F (1, 292) 152,28, $p < .001$, and a significant treatment by trials interaction, F (2, 292) 4.62, $p < .01$. The means are shown in Figure 22.

The significant treatment by trials interaction is explanied by the fact that the Blend and Traditional treatments both made greater gains than did the Thematic group. Another way of

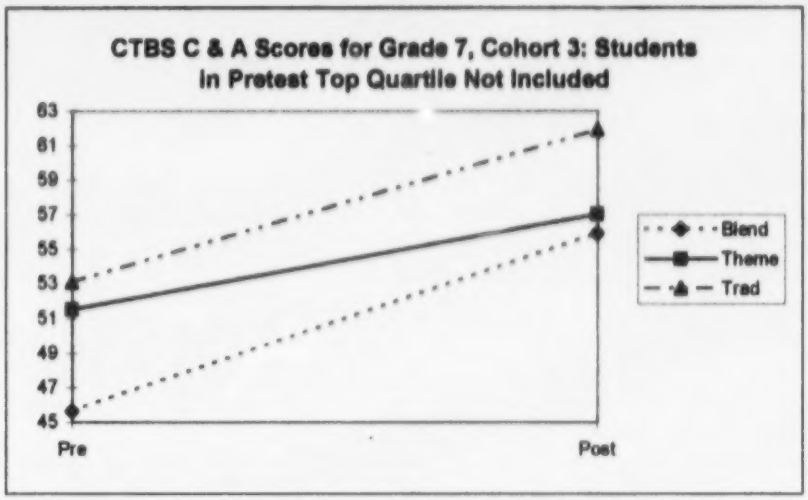

Figure 22

seeing this is by means of an analysis of covariance, with pretest scores as the covariate. The effect for condition was significant, F (2, 291) 3.26, $p < .05$. The adjusted means of the Blend (58.97) and Traditional (58.99) treatments did not differ, but both were significantly greater than the adjusted mean for the Thematic treatment (55.38). Again, we would point out that the success of the Blend group appears to be due in large measure to classes taught by a teacher who taught in the Thematic condition during the prior two years, and who continued to use and expand much of the curriculum she had used when she taught Theme classes.

We conducted the same kinds of analysis on data for the 19945-96 eighth graders, eliminating from the analyses all students who scored in the top quartile on the pretest. The main effects for treatment, F (2, 202) 15.55, $p < .001$, and trials, F (1,202) 15.84, $p < .001$ were both significant, as was the treatment by trials interaction, F (2, 202) 9.78, $p < .001$. The interaction is depicted in Figure 23.

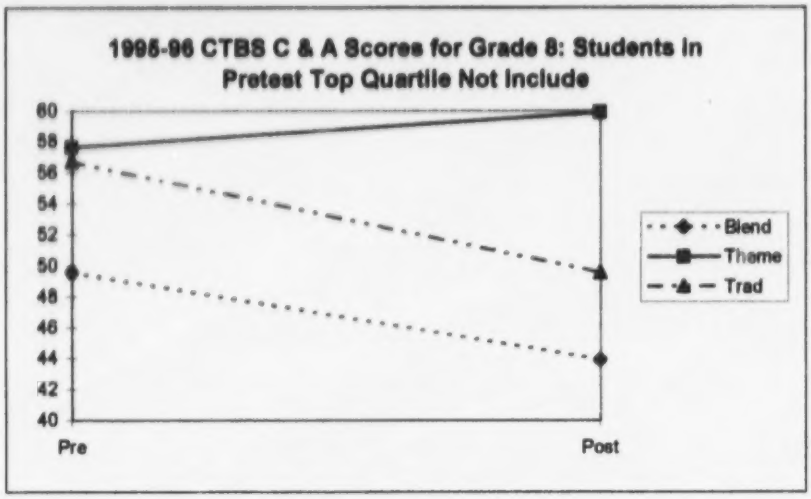

Figure 23

As can be seen in Figure 23, the significant interaction is explained by the fact that the Thematic group made gains whereas scores declined from pre- to posttesting in both the Traditional and Blend treatments. However, the Theme group's gain, t (df 54) = 1.39, *ns*, did not attain significance, whereas the losses of the Traditional, t (df 60) = 4.82, $p < .001$ and Blend, t (df 88) = 4.06, $p < .001$ groups were statistically significant, These results are parallel to those of the original analyses which included all students for whom pre and posttest scores were available. From these analyses we can conclude that "topping out" among students in the upper quartile did not affect the overall pattern of achievement results.

The teachers expressed pleasure with the outcomes of the project, with the working relationships developed with the business and university partners, and with the positive attitudes that have evolved from the partnership. The teachers all feel that they have witnessed highly beneficial effects, not only from the development of the Mathematics in Manufacturing curriculum and instructional materials, but also from the job shadowing opportunities their students experienced at the Radionics facility and the shadowing of university students at UCSC. They assert that

these support systems are just as important as changes in curricular and pedagogical approaches to mathematics education, and they express hope that ways and means can be found to sustain these activities.

Contrasting with their positive feelings about the CAPP project and what it enabled them to accomplish is a perception that the mathematics faculty is misperceived and unappreciated by district administrators and by some faculty at the site as well. As a group, the mathematics teachers appear to see themselves as proactive, concerned educators who have made significant changes to improve the success of their students and their access to a rich educational experience. They are distressed that others at the site and in the district do not seem to realize that. The perception of those of us from the University is that there are substantial variations among teachers in the degree to which they have changed their instructional approaches, but that all have changed in some degree, and that, in general, some very major instructional improvements have been accomplished. Certainly, this is not a group that has jumped on the mathematics reform bandwagon in a vocal and uncritical way. We speculate that their preference for results over rhetoric may contribute to external perceptions of their willingness to transform their curriculum and instructional practices. Nevertheless, in varying degrees they have tried new practices and been willing to revise their practices when change seemed warranted by their own experience and by the evaluation data. But they do trust their own experience at least as much as they trust theory or data from testing. This is evident in the fact that, in spite of favorable CTBS data, there is still little support for heterogeneous grouping, at least in the form it was implemented within the scheduling restrictions imposed by the existence of the GATE program.

References

Cocking, R. R., and Chipman, S. (1988). Conceptual issues related to mathematics achievement of language minority children. In R. R. Cocking and J. P. Mestre, (Eds.), *Linguistic and cultural influences on learning mathematics.* (pp. 17-46). Hillsdale, NJ: Erlbaum.

Chavez, S. (1993, Sept. 3) L.A. schools' bilingual program failing, state says. *Los Angeles Times*, v112:A1, col 5.

Dossey, J. A., Mulls, I. V. S., Lindquist, . M., and Chambers, D. L. (1988). *The mathematics report card: Are we measuring up*? Princeton, NJ: Educational Testing Service.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.

Dweck, C. S. and Leggett, E. L. (1988). A socio-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.

Elliot, E., and Dweck, C. S. (1988). Goals: an approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*, 5-12.

Haladyna, T., Shaughnessy, J., and Shaughnessy, J. M. 1983. A causal analysis of attitude toward mathematics. *Journal for Research in Mathematics Education*, 14, 19-29.

Harter, S. (1985). *Manual for the self-perception profile for children.* Denver: University of Denver.

Haycock, K., and Navarro, M. S. (1988). *Unfinished business: Fulfilling our children's promise.* Oakland, CA: The Achievement Council.

Henderson, R. W. (1991). Development of a measure of motivational goal orientations for mathematics learning. Santa Cruz: Center for Research on Cultural Diversity and Second Language Learning, University of California.

Henderson, R. W. (1995, April). Middle school mathematics for students of Mexican descent: A thematic approach to contextualization of instruction. National Center for Research on Cultural Diversity and Second Language Learning, University of California, Santa Cruz, CA. Eric Document Reproduction Service No. ED 386 368.

Henderson, R. W., St. John, L., & Youpa, D. (1994, Nov.). *Internal evaluation of a thematic approach to mathematics in the middle school: A school-university-business partnership (33 pp).* Unpublished technical report submitted to the California Academic Partnership Program. Santa Cruz: The University of California.

Henderson, R. W., & Landesman, E. M. (1995). Attitudes, motivational goal structures, and achievement in mathematics among middle school students of Mexican descent: Effects of thematically integrated instruction. *Journal of Educational Research, 88*, 290-300.

Kamm, T. (1994, April 20). Brizil swiftly becomes major auto producer as trade policy shifts: VW, Ford, GM are building new low-priced models for eager populace. *The Wall Street Journal, Pp. A1, A4.*

LaFranchi, H. (Dec 28, 1984 LaFranchi, Howard. (1984, Dec. 28). The kids we can't afford to waste; urgent need: better education America's fastest-growing minority. *Christian Science Monitor.* v77, 19, col 2.

MacCorquodale, P. (1988). Mexican-American women and mathematics: Participation, aspirations, and achievement. In R. R. Cocking and J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics*, (pp. 137-160). Hillsdale, NJ: Erlbaum.

McLeod, D. B. (1985). Affective issues in research on teaching mathematical problem solving. In *Teaching and learning mathematical problem solving: Multiple research perspectives*, ed. E. A. Silver, (pp. 267-279). Hillsdale, NJ: Lawrence Erlbaum.

Mendoza, M. (1995, Jan. 15). Together, not equal: Despite millions of dollars spent at Pajaro Valley schools, achievement lags, especially for Hispanic students. *Santa Cruz County Sentinel.* pp. A1, A10.

Moss, C. R. (1995). An exploration of the effects of assessment on the mathematical writings of seventh grade students. Unpublished Master's, University of California, Santa Cruz.

O'Donnell, S. M. (1995). Alternative assessment in mathematics: The effect of performance-based assessment on the achievement of seventh grade students. Unpublished Master's Thesis, University of California, Santa Cruz.

Peterson, P. L. (1988). Teaching for higher-order thinking in mathematics: The challenge for the next decade. In D. A. Grouws, T. J. Cooney and D. Jones (Eds.), Perspectives on research on effective mathematics teaching, Vol. 1, (pp. 2-26). Reston VA: National Council of Teachers of Mathematics.

Policy Analysis for California Education. (1990). *Conditions of education in California.* Berkeley: Policy Analysis for California.

Regional Math Network. (1987). *Quincy market: A problem-solving unit from the Regional Math Network.* Palo Alto: Dale Seymour Publications.

Reich, R. B. (1991). *The work of nations: Preparing ourselves for 21st century capitalism.* New York: Alfred A. Knopf.

Romberg, T. A., & Zarinnia, A. (1987). Consequences of the new world view to assessment of students' knowledge of mathematics. In T. A. Romberg & D. M. Stewart (Eds.), *The monitoring of school mathematics: Background papers: vol. 2. Implications from psychology: Outcomes of instruction* (pp. 153-201). Madison: Wisconsin Center for Education Research, University of Wisconsin -- Madison.

St. John, L. (1996). The social construction of motivation: Small group work among Latino and Euro-American students in eighth grade mathematics classrooms. Unpublished doctoral dissertation, University of California, Santa Cruz.

Silver, E. A. (1985). Research on Teaching Mathematical Problem Solving: Some Underrepresented Themes and Needed Directions. In Edward A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives*, (pp. 247-266). Hillsdale, NJ: Lawrence Erlbaum.

Sowder, J. T. (Ed.). (1989). *Setting a research agenda. Research Agenda for Mathematics , Vol. 5.* Reston, VA: The National Council of Teachers of Mathematics.

Tharp, Roland G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context.* New York: Cambridge University Press.

Valencia, R. R. (1991). The plight of Chicano students: An overview of schooling conditions and outcomes. In Richard R. Valencia (Ed.), *Chicano School Failure and Success: Research and Policy Agendas for the 1990s*, (pp. 3-26). London: Falmer Press.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Vygotsky, L. S. (1978). *Mind in society.* (M. Cole, V. John-Steiner, S. Schribner, & E. Saberman, Eds.). Cambridge, MA: Harvard University Press.

Zarinnia, E. A., & Romberg, T. A. (1987). A new world view and its impact on school mathematics. In T. A. Romberg & D. M. Stewart (Eds.), *The monitoring of school mathematics: Background papers: Vol. 1. The monitoring project and mathematics curriculum*, pp. 21-61. Madison: Wisconsin Center for Education Research, University of Wisconsin -- Madison.

ENDNOTE

1. We want to express our appreciation to Sally Parker of the Salinas Unified High School District for her assistance in organizing the data collection effort and obtaining testing and demographic data from school district computer files. We are also grateful to the teachers and administrators of Washington Middle School for their cooperation with the data collection effort, and to the teachers and students whose participation made the effort on which this evaluation is based possible. We are especially indebted to project co-directors Margaret White of the Salinas Unified High School District, and Edward M. Landesman, Research Professor of Mathematics at the University of California, Santa Cruz, for their dedicated commitment running a project of the highest quality. We also thank Daniel Youpa for his assistance with data management and analysis.